WGET

GNU Wget es un recuperador por red no interactivo.
Modo de empleo: wget [OPCIÓN]... [URL]...

EJEMPLO:

- `wget -m -k -Pweb http://eos.cnice.mecd.es/mem2000/algebra/index.html`
  donde:
  - **`-m` significa que queremos realizar un mirror (copia exacta del sitio, con sus carpetas, subcarpetas, ...) y es un atajo equivalente a -r -N -l inf -nr.**
  - **`-k` convierte los enlaces no relativos en relativos.**
  - **`-Pweb`** *nos creará el subdirectorio web, donde se almacenará todo lo descargado.*
  - **`http://...`** *es la dirección de la página inicial desde la que comenzará la descarga, siguiendo los enlaces existentes en ella. También podemos poner poner la dirección de una carpeta sin página de inicio.*

Y por si te interesa conocer más opciones a continuación tienes el todas las del "man" en castellano.

OPCIONES:

Inicio:		
-V,	**--version**	muestra la versión de wget y termina.
-h,	**--help**	muestra esta ayuda.
-b,	**--background**	pasa a segundo plano al iniciar.
-e,	**--execute=ORDEN**	ejecuta una orden como las de `.wgetrc'.
Fichero de entrada y registro:		
-o,	**--output-file=FICHERO**	registra los mensajes en FICHERO.
-a,	**--append-output=FICHERO**	añade los mensajes a FICHERO.
-d,	**--debug**	imprime la salida de depurado.
-q,	**--quiet**	modo silencioso (no muestra ninguna salida).
-v,	**--verbose**	modo informativo (predeterminado).
-nv,	**--non-verbose**	muestra el mínimo necesario de información.
-i,	**--input-file=FICHERO**	descarga las URLs que haya en FICHERO.
-F,	**--force-html**	trata el fichero de entrada como HTML.
-B,	**--base=URL**	añade URL delante de los enlaces relativos en el fichero -F -i.
	--sslcertfile=FICHERO	certificado opcional del cliente.
	--sslcertkey=FICHERO	llave opcional para este certificado.
	--egd-file=FICHERO	fichero del socket EGD.
Descarga:		
	--bind-address=DIRECCIÓN	realiza un bind a la DIRECCIÓN (máquina o IP) en la máquina local.
-t,	**--tries=NÚMERO**	establece en NÚMERO el número de reintentos(0 no pone límite).
-O,	**--output-document=FICHERO**	escribe los documentos en FICHERO.
-nc,	**--no-clobber**	no sobreescribir ficheros existentes. o utilizar sufijos .#
-c,	**--continue**	continuar recuperando un fichero existente.
	--dot-style=ESTILO	establece el estilo de la pantalla de recuperación.
-N,	**--timestamping**	no recupera ficheros más viejos que los locales.
-S,	**--server-response**	imprime la respuesta del servidor.
	--spider	no recupera nada.
-T,	**--timeout=SEGUNDOS**	establece el tiempo de espera de lectura en SEGUNDOS.
-w,	**--wait=SEGUNDOS**	espera SEGUNDOS entre recuperaciones.
	--waitretry=SEGUNDOS	espera 1...SEGUNDOS entre reintentos.
	--random-wait	espera de 0 a 2*WAIT segundos entre reintentos.
-Y,	**--proxy=on/off**	habilita/deshabilita el uso de proxies.
-Q,	**--quota=NÚMERO**	establece la cuota de recuperación en NÚMERO.
	--limit-rate=TASA	limita la tasa de descarga a TASA.

Directorios:

-nd	**--no-directories**	no crea directorios.
-x	**--force-directories**	fuerza la creación de directorios.
-nH,	**--no-host-directories**	no crea directorios en el anfitrión
-P,	**--directory-prefix=PREFIJO**	guarda ficheros en PREFIJO/...
	--cut-dirs=NÚMERO	descarta NÚMERO componentes del directorio remoto.

Opciones de HTTP:

	--http-user=USUARIO	establece que el usuario de http es USUARIO.
	--http-passwd=CLAVE	utiliza CLAVE como contraseña de http.
-C,	**--cache=on/off**	(des)habilita la caché del servidor de datos (normalmente
-E,	**--html-extension**	guarda todos los ficheros de texto/html con la extensión .html.
	--ignore-length	ignora el campo 'Content-Length' de la cabecera.
	--header=TEXTO	inserta el TEXTO entre las cabeceras.
	--proxy-user=USUARIO	establece que el usuario del proxy es USUARIO.
	--proxy-passwd=CLAVE	utiliza CLAVE como contraseña del proxy.
	--referer=URL	incluir cabecera 'Referer: URL' en petición HTTP.
-s,	**--save-headers**	guarda las cabeceras de HTTP en un fichero.
-U,	**--user-agent=AGENTE**	identificarse como AGENTE en vez de Wget/VERSIÓN.
	--no-http-keep-alive	deshabilita las conexiones persistentes de HTTP.
	--cookies=off	no utiliza cookies.
	--load-cookies=FICH.	carga las cookies desde FICH. antes de la sesión.
	--save-cookies=FICH.	guarda las cookies en FICH. tras la sesión.

Opciones de FTP:

-nr	**--dont-remove-listing**	no elimina los ficheros '.listing'.
-g,	**--glob=on/off**	habilita/deshabilita el uso de comodines en ficheros.
	--passive-ftp	usa el método de trasferencia "pasivo".
	--retr-symlinks	al descender por los directorios, recupera los ficheros a los que

Recuperación recursiva:

-r,	**--recursive**	succionado de web recursivo, ¡úselo con cuidado!
-l,	**--level=NÚMERO**	profundidad máxima de recursión (0=ilimitada)
	--delete-after	borra los ficheros locales una vez descargados.
-k,	**--convert-links**	convierte los enlaces no relativos en relativos.
-K,	**--backup-converted**	antes de transformar el fichero X, hace una copia como X.orig.
-m,	**--mirror**	atajo equivalente a -r -N -l inf -nr.
-p,	**--page-requisites**	descarga todas las imágenes, etc. necesarias para mostrar la página

Aceptar/rechazar recursivo:

-A,	**--accept=LISTA**	lista, separada por comas, de extensiones aceptadas.
-R,	**--reject=LISTA**	lista, separada por comas, de extensiones rechazadas.
-D,	**--domains=LISTA**	lista, separada por comas, de dominios aceptados.
	--exclude-domains=LISTA	lista de dominios rechazados, separados por comas.
	--follow-ftp	sigue enlaces a FTP desde documentos HTML.
	--follow-tags=LISTA	lista de etiquetas HTML a las que se sigue.
-G,	**--ignore-tags=LISTA**	lista de etiquetas HTML que se ignorarán.
-H,	**--span-hosts**	ir a anfitriones distintos en la recursión.
-L,	**--relative**	sigue sólo enlaces relativos.
-I,	**--include-directories=LISTA**	lista de directorios permitidos.
-X,	**--exclude-directories=LISTA**	lista de directorios excluidos.
-np,	**--no-parent**	no asciende al directorio padre.

NOTA: Los argumentos obligatorios para las opciones largas son también obligatorios para las opciones cortas.

José Sánchez Muñoz
Asesor de Nuevas Tecnologías
Centro de Profesores y de Recursos de Plasencia
jose@josesanchez.net
http://www.josesanchez.net